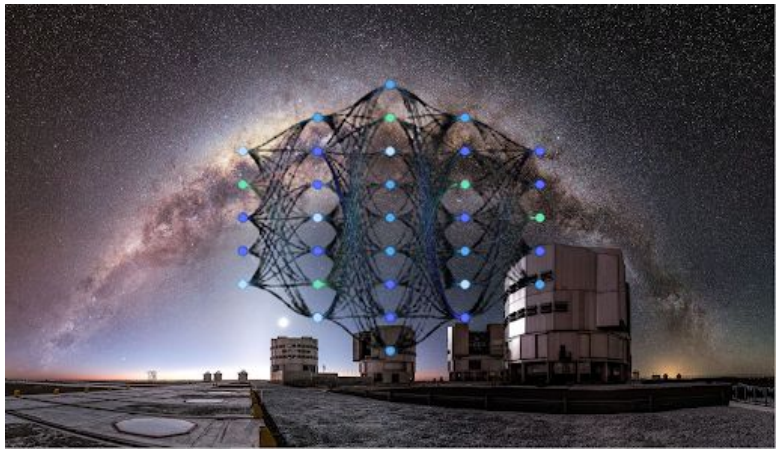
Target Selection for a redshift limited survey with Machine Learning

Gursharanjit Kaur

Maciej Bilicki and WAVES photo-z group

Center for Theoretical Physics, PAN, Warsaw



July 10, 2024

MACHINE LEARNING FOR ASTROPHYSICS
2ND EDITION CATANIA, 8-12 JULY, 2024

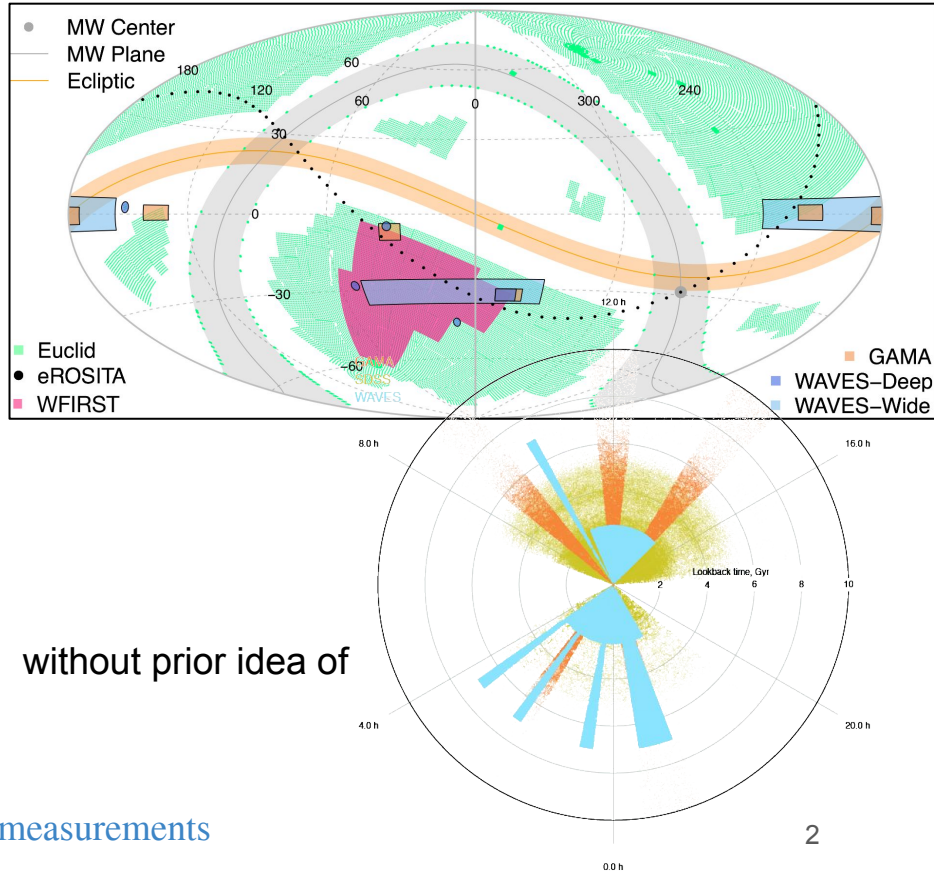
WAVES
Wide Area Vista Extragalactic Survey

4MOST (4 meter Multi Object Spectroscopic Telescope)
Spectrograph to be hosted at VISTA telescope

WAVES: Galaxy Evolution Survey with 4MOST spectrograph
In two sub-surveys
WIDE
0.9 million galaxies
Z band (central wavelength 0.88 μm)
magnitude , **ZMag ≲ 21.1, redshift z ≲ 0.2**

Probing significantly lower galaxy and halo masses in the low-redshift than before e.g GAMA
Survey Success Criteria: **Completeness of 0.95**
WIDE has complete overlap with KiDS (Imaging survey)
Input Photometry is from KiDS+VIKING

Driver et al, 2019



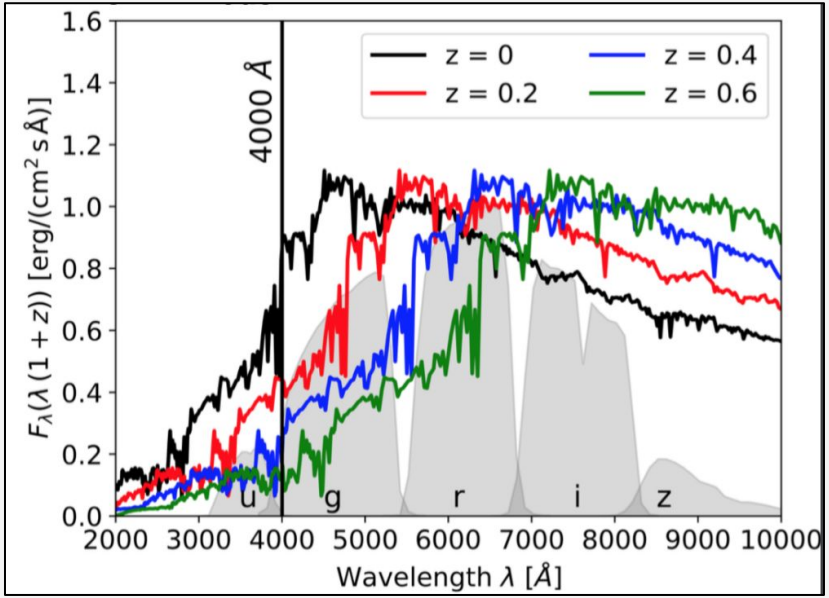
Challenge:

Selecting the targets for spectroscopic redshift estimation without prior idea of redshift!

General Solution:

Photometric redshift estimation using the color and magnitude measurements

Photometric redshift estimation



Credits: Ben Hoyle Talk, Munich, 2018 SDSS filters

Basic: Identify the strong continuum features (Balmer, Lyman break) in spectra and multiband data

★ **Template fitting approach**
- Synthetic photometry templates based on complex galaxy evolution models
- Real data compared to synthetic photometry
- Large range of template spectra and redshift
- Some Codes: TopZ, LEPHARE, HYPERZ

★ **Non-parametric/ Machine Learning**
- Data with photometry and spectroscopic redshift as training set
- Minimize the difference between spec_z and ML_z
- Using Neural Networks, Self Organizing Maps(SOM)

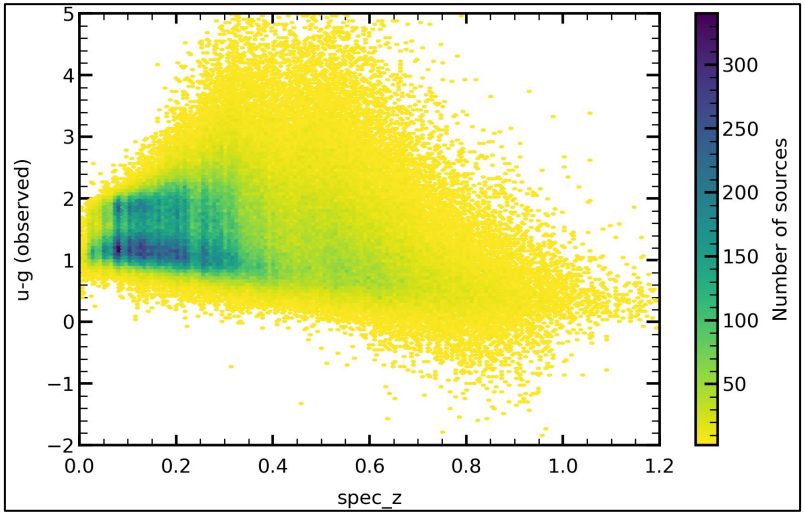
For the target selection for 4MOST WAVES-WIDE,

- Instead of redshift estimation, we need galaxies to be within a redshift and a magnitude limit
- **A classification problem instead of a regression problem**: either **redshift ≤ 0.2 or redshift > 0.2**
- **Using ML Classification (XGBoost), we get a probability of z < z_lim**

Target selection catalog

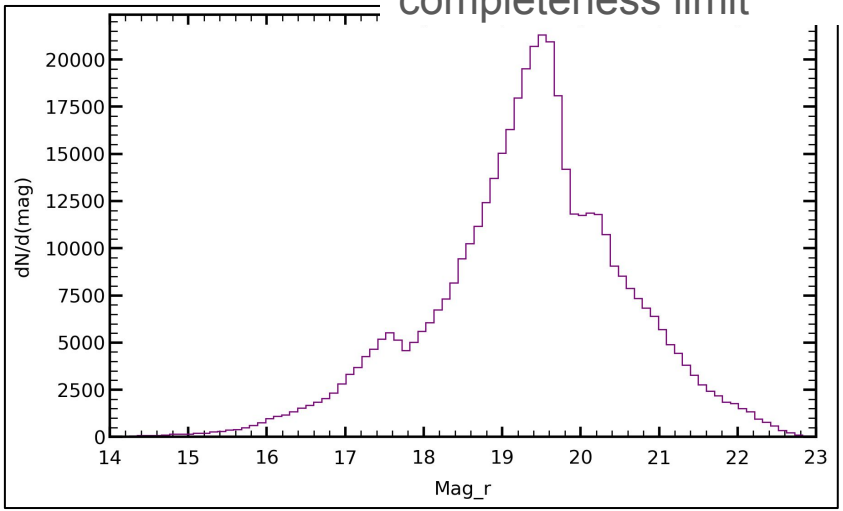
- Imaging from VIKING (Z-band selection) and KiDS (for photo-zs): **ugriZYJHK** (0.3 μm to 2.2 μm)
- PROFOUND photometry (total galaxy flux) (Robotham et al. 2018)
- with corresponding spectroscopic redshift from cross matching surveys (GAMA, COSMOS, DESI, …)
- Incompleteness at fainter end is due to bias of spectroscopic survey towards Luminous Red Galaxies (LRGs), ELGs

KiDS VIKING reprocessed photometry



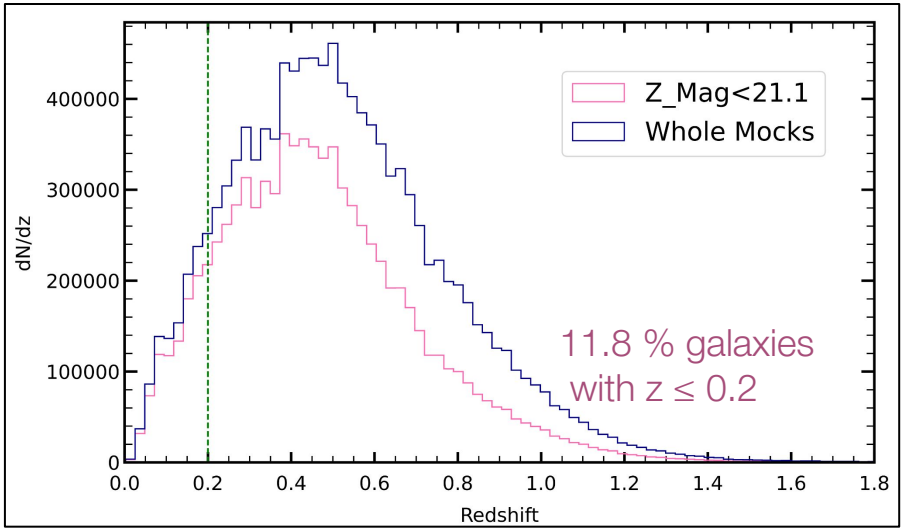
Redshift from spectroscopic compilation

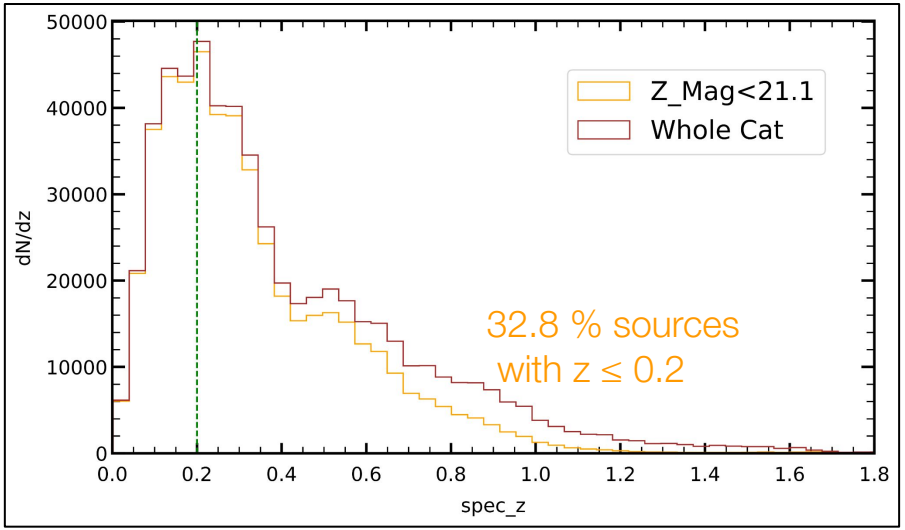
GAMA completeness limit



Target selection catalog

Spectroscopic sample from redshift from cross matching surveys (SDSS, GAMA, COSMOS, DESI, …)

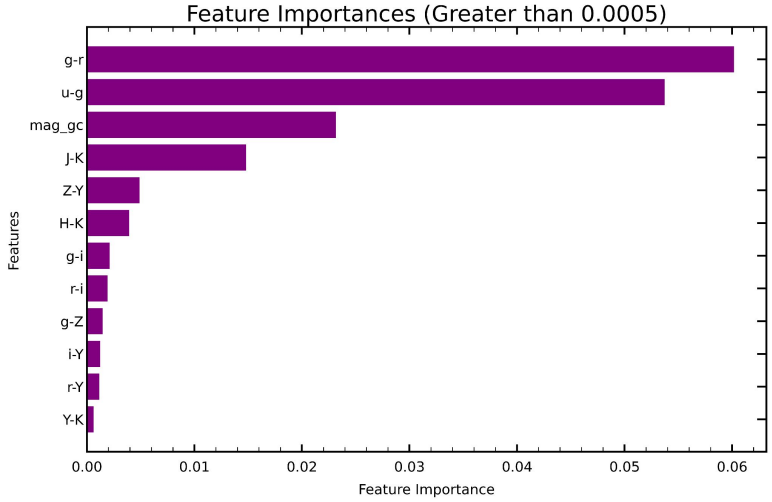


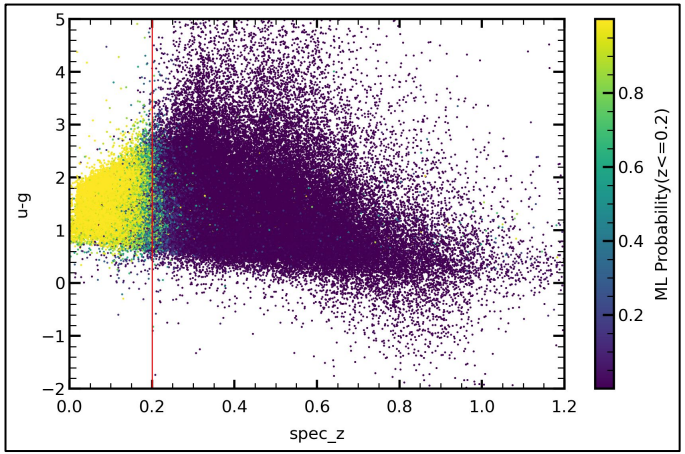


Redshift distribution from **SHARK** mocks

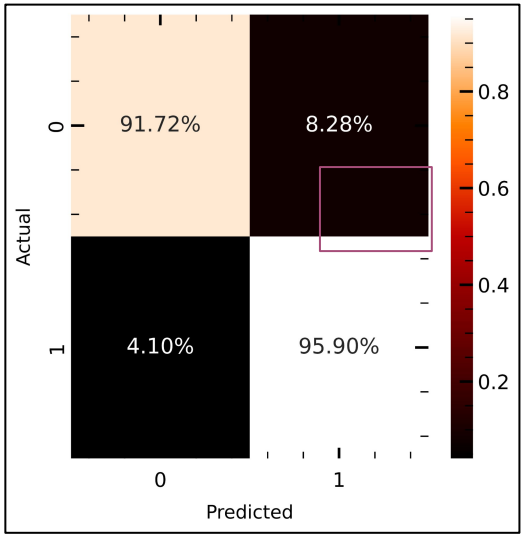
Redshift distribution from **Selection catalog**

Results





Using Magnitudes,Colors as features



Purity:
91.58%
Completeness:
91.72%

- ★ A photometry based redshift classification pipeline
- ★ Can provide the probability of the target lying within survey target limit
- ★ Based on the survey preferences, the probability threshold can be varied to achieve either higher purity or completeness